Unifying AoI Minimization and Remote Estimation — Optimal Sensor/Controller Coordination with Random Two-way Delay

Cho-Hsin Tsai and Chih-Chun Wang
School of Electrical and Computer Engineering, Purdue University, USA
Email: {chtsai, chihw}@purdue.edu

***Abstract*—The ubiquitous usage of communication networks in modern sensing and control applications has kindled new interests on the timing-based coordination between sensors and controllers, i.e., how to use the "waiting time" to improve the system performance. Contrary to the common belief that a zero-wait policy is always optimal, Sun *et al.* showed that a controller can strictly improve the data freshness, the so-called Age-of-Information (AoI), by postponing transmission in order to lengthen the duration of staying in a good state. The optimal waiting policy for the *sensor* side was later characterized in the context of remote estimation. Instead of focusing on the sensor and controller sides separately, this work develops the optimal joint sensor/controller waiting policy in a Wiener-process system. The results can be viewed as strict generalization of the above two important results in the sense that not only do we consider joint sensor/controller designs (as opposed to sensor-only or controller-only schemes), but we also assume random delay in both the forward and feedback directions (as opposed to random delay in only one direction). In addition to provable optimality, extensive simulation is used to verify the performance of the proposed scheme in various settings.**

I. Introduction

The omnipresence of a massive number of portable devices has led to increasing focus on systems with multiple sensors and controllers interconnected by wireless/wired communication networks. Many new research directions have been initiated, including healthcare, energy management systems, cloud data infrastructure (see [1]–[5]). In this work, we study the question: How to optimally coordinate the sensor and the controller when there is random delay in both the forward and backward directions? We begin the analysis by observing there are two distinct ways of timing-based system optimization: *data-freshness control* and *state-based sampling*.

Data-freshness control: In this way, the controller is the one who actively maintains the data-freshness of the system. For example, say the goal is to lower the risk of heart attacks of the patients. One way to achieve the aim is for the hospital (controller) to make sure that the blood pressure (BP) or the heart rate (HR) records of the patients are as fresh as possible. To this end, the hospital should intermittently request the patients (sensors) to measure their latest BP or HR and send in the reports. In practice, any sensor-to-controller measurement packet inevitably experiences some delay and is thus always "stale" to some degree. The controller (hospital) must decide how to optimize its request schedule in order to maximize the data-freshness of its records.

One breakthrough of the data-freshness control is the introduction of a new metric, Age-of-Information (AoI) [6], the corresponding minimization algorithms [7], and its numerous follow-up results [8]–[11]. For instance, a "generate-at-will" model was studied in [12], [13], which has the potential of considerable energy savings.

In general, AoI minimization behaves quite differently from throughput optimization. For example, the zero-wait policy [14] was proven to be throughput optimal. In contrast, in a recent important work [15], Sun *et al.* proved that the zero-wait policy is far from optimal in terms of the average AoI. Sun *et al.* also characterized the optimal "waiting time" that can provably minimize the average AoI.

State-based sampling: Unlike the data-freshness control, in this line it is the sensor that actively optimizes the overall system.[1] Continue from the aforementioned example of lowering the risk of heart attacks. The state-based sampling approach is for the patient (sensor) to measure his/her own BP and HR continuously and report it when and only when the BP/HR shows elevated risk. Once the hospital (controller) receives the report, some treatment (action) is prescribed to bring the BP/HR back to normal. The patient will not send in any new report unless the BP/HR once again exhibits new concerns.

This work was supported in parts by NSF under Grant CCF-1422997, Grant CCF-1618475, and Grant CCF-1816013.

[1]The best way to determine whether a scheme is *controller-based* or *sensor-based* is to examine in which physical location the decision is made, since their distinct locations naturally lead to different roles and, more importantly, *asymmetric access to the underlying random states and timing information*. Also see our discussion in Sec. II. However, such a definition does not apply to many existing results. The reason is that with the assumption of instantaneous ACK feedback, one node has full access to the information available at the other node, which breaks the information asymmetry and thus blends the roles of sensors and controllers. The second way of classification is to see whether the algorithm has instantaneous access to the (random) value of the measurement and explicitly uses the measurement to decide whether to transmit or not. If so, it is a senor-side algorithm, e.g., the remote estimation scheme in [16]. Otherwise, it is a controller-side algorithm. Under this methodology, the AoI minimization scheme in [15] is classified as controller-based even though it is actually executed by the sensor. In this particular example, one can envision "the controller" as a separate computer program within the physical sensor that tells the sensor when to transmit *without using the actual measurement data*.

The focus of this direction is thus to design schemes that detect the changes in signal values and opportunistically send the updates when the need arises. This direction is generally termed the (state-based) sampling schemes for remote estimation. An early work [17] showed that a *threshold policy* can lower the estimation error. Later it was shown that the *threshold policy* is optimal for a variety of settings, including cellular networks [18], noisy channels [19] and multi-dimensional Wiener processes in [20]. In [16], Sun *et al.* generalized the setting of [20] by adding a queue with random service time between the sampler and estimator, and showed that the threshold policy is optimal. Further discussion of the threshold policy will be provided in Sec. IV.

The main motivation of this work is two-fold. Firstly, since both directions can significantly improve the system performance, one cannot help but wonder whether the controller (hospital) and the sensor (patient) can work jointly and attain global optimality. Secondly, since we are interested in remote systems with non-collocated sensors and controllers, there is likely to be random delay for both the sensor-to-controller and the controller-to-sensor directions. Nonetheless, the existing results [15] and [16] and all the aforementioned works assume random delay in only one direction, plus idealized zero-delay acknowledgement (ACK) for the other direction. It is thus of paramount interest to study how the existing schemes perform under a more practical 2-way delay model and see whether one can further improve the performance by optimally accounting for the 2-way delay.

Our key contributions are summarized as follows.

(i) We propose a new framework that unifies the controller-side AoI minimization problem [15] and the sensor-side remote estimation problem [16].

(ii) Our framework allows for arbitrary random 2-way delay distributions, does not rely on idealized instantaneous ACK, and thus would be more suitable for practical applications where random delay is universally present in both directions.

(iii) We derive the optimal joint sensor/controller policy under the proposed new setting. For comparison, existing works focus either on the sensor or on the controller and take into account random delay in only one direction. The double relaxation from a single-node policy to a joint policy and from the 1-way delay to 2-way delay represents an outstanding advancement over the state of the art.

(iv) Our findings also show that blindly applying the existing schemes [15], [16] to practical systems with 2-way random delay could lead to significant performance loss and the results can sometimes be much worse than a naive zero-wait policy.

The remainder of the paper is organized as follows. In Sec. II, our detailed system model and problem formulation are presented. Our main results are outlined in Sections III and IV with the sketches of the proofs provided in Sec. V. Numerical results are reported in Sec.VII. We conclude our work in Sec. VIII.

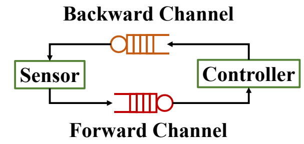


Fig. 1: A sensor/controller system with 2-way delay.

II. Model and Formulation

A. System Model

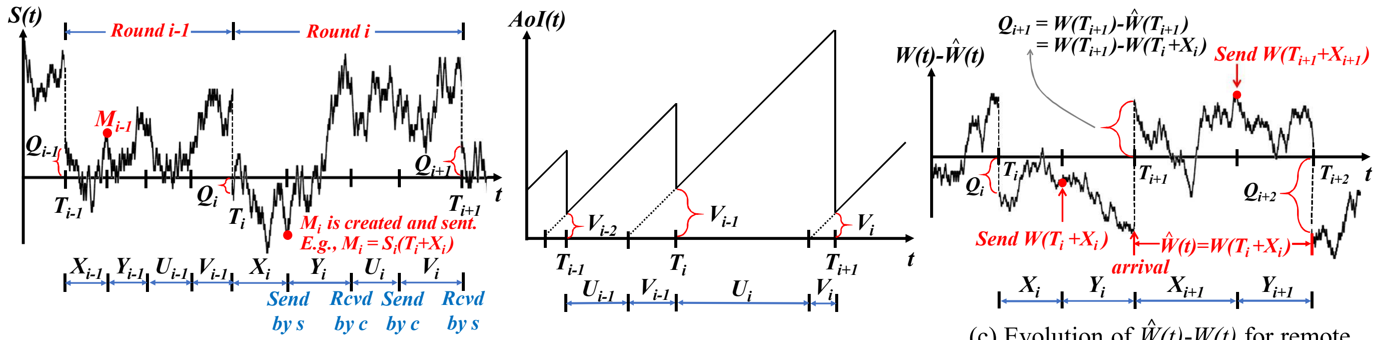
Our system model is best depicted in Fig. 1, which consists of a sensor, a controller, a forward sensor-to-controller channel and a backward controller-to-sensor channel. It is worth noting that we use the terms of sensor and controller in their broadest sense. The sensor node is not limited to a physical sensor that measures the location/temperature of the environment. Instead, it can be any data-generating node, including a database server, a video-streaming source, etc. Also, the controller is not restricted to a node directly commanding an actuator. Instead, it can be any decision making component, e.g., computation of the inferred status of the remote database, or the video processing applications that render the actual video. Further discussion about the roles of each node can be found in Sec. II-C and II-D.

Each of the two channels incurs random transmission delay. With two-way random delay in the communication loop, the timing information at the sensor and the controller is inherently unsynchronized. Specifically, the waiting policy of the sensor (resp. controller) does not have instantaneous access to the status of the underlying network and has to wait for the delayed response from the controller (resp. sensor). *This two-way delay model and the resulting double time asynchrony where neither the sensor nor the controller has the perfect global timing information is the most distinguishing feature of this work.* For comparison, most existing works [7]–[11], [15], [16], [21], [22] assume one node has perfect network-wide timing information, which may not hold in practice where random delay is universally present.

We now explain our system model. We denote system state as $S(t)$, for which we shift the values so that the origin $S(t) = 0$ is the most desired system state. The value of $S(t)$ may drift away from zero as time proceeds. We assume the evolution of $S(t)$ is related to a Wiener process $W(t)$ [23], [24], a widely used (though idealized) model of system states.[2] The detailed system evolution is defined as follows, and the corresponding illustration is provided in Fig. 2a.

Time sequences: The system consists of four discrete-time real-valued random processes X_i, Y_i, U_i, and V_i for all i. X_i is the i-th waiting time at the sensor; Y_i is the random delay for the i-th use of the sensor-to-controller channel; U_i is the i-th waiting time at the controller; V_i is the random delay for the i-th use of the controller-to-sensor channel.

[2] Some applications of the Wiener process model include unmanned aerial vehicles (UAVs) [25], biosensing schemes [26] and mobile networks [27].



(a) Our setting. (b) Evolution of AoI [15]. (c) Evolution of $\hat{W}(t)$-$W(t)$ for remote estimation [16].

Fig. 2: Illustration of system evolution for different problem formulations.

The values of $\{X_i\}$ to $\{V_i\}$ can be used to derive another time sequence $\{T_i\}$ as follows: $T_1 \triangleq 0$ and $T_{i+1} \triangleq T_i + X_i + Y_i + U_i + V_i$ for all i. We call the interval $[T_i, T_{i+1})$ as the *i-th round,* which consists of the i-th waiting time of the sensor, the delay of the i-th use of the forward channel, the i-th waiting time of the controller, and the delay of the i-th use of the backward channel. Clearly, the value T_i is the beginning of the i-th round.

We now describe the system behavior in the i-th round.

Reset-to-Q_i at the sensor: At time T_i, the sensor has received the message from the controller in the previous $(i-1)$-th round. It is very convenient to view the message as a *reset command.* We assume that upon receiving the reset command, the system state at time T_i will be reset to a random value Q_i, which is the (random) *initial value* of the i-th round. For example, with thermal noise it may be impossible to set the state value to be exactly 0 with infinite accuracy. The value Q_i thus models the residual randomness after reset, if there is any. We assume $\{Q_i\}$ is i.i.d. with $\mathbb{E}\{Q_i\} = 0$.

After reset-to-Q_i, the system state will evolve according to a Wiener process $W(t)$ during the entire i-th round, until it is once again reset to Q_{i+1} at time T_{i+1}. The state value in the i-th round, denoted by $S_i(t)$, is captured by

$$S_i(t) = W(t) - W(T_i) + Q_i, \text{ for } t \in [T_i, T_{i+1}). \quad (1)$$

We sometimes drop the subscript i and simply use $S(t)$.

Waiting time at the sensor: The sensor has the ability of waiting for an arbitrary amount of time $X_i \geq 0$, which is the so-called "generate-at-will" model in [12], [13], [15]. The random variable X_i is a *stopping time* with respect to the filtration generated by $\{S(\tau) : \tau \leq t\}$. That is, the sensor observes the evolution of the system states and causally decides when to stop waiting and start transmission.

Upon transmission, the sensor sends (T_i, X_i, M_i) to the controller, where T_i is the time it receives "reset" in the previous round; X_i is the waiting time; M_i is the additional message, which is again decided by the past system states.

Random delay in the forward direction: The tuple (T_i, X_i, M_i) sent by the sensor at time $T_i + X_i$ will arrive at the controller at time $T_i + X_i + Y_i$. The transmission delay Y_i is i.i.d. and is independent from the rest of the system.

Waiting time at the controller: Since the message is time-stamped (containing (T_i, X_i)), the controller can infer the value of the forward transmission delay Y_i by subtracting $T_i + X_i$ from the actual arrival time $T_i + X_i + Y_i$. The waiting time U_i at the controller is then a function of all the previous messages and timing information $\{(T_j, X_j, Y_j, M_j) : j \leq i\}$.

Random delay in the backward direction: At time $T_i + X_i + Y_i + U_i$, the controller sends a reset signal, which experiences a random delay V_i and will reach the sensor at time $T_{i+1} \triangleq T_i + X_i + Y_i + U_i + V_i$. The $(i+1)$-th round then begins, and we go back to reset-to-Q_{i+1} at the sensor. Again, we assume V_i is i.i.d. and is independent from the rest of the system.

Technical assumptions: Similar to [15], [16], we assume the statistics of $\{Q_i\}$, $\{Y_i\}$, and $\{V_i\}$ are known to both the sensor and the controller and $0 < \mathbb{E}\{Y_i\} + \mathbb{E}\{V_i\} < \infty$, and $\mathrm{Var}\{Q_i\} + \mathrm{Var}\{Y_i\} + \mathrm{Var}\{V_i\} < \infty$.

Final remark: The non-negativity $X_i \geq 0$ (resp. $U_i \geq 0$) prohibits the sensor (resp. controller) to transmit before receiving the ACK, as in most TCP-based control protocols [28]. It is possible to design an even better scheme that transmits anticipatively before receiving any ACK, which, however, is beyond the scope of this work.

B. The Objective

For any given scheme $\{X_i\}$ and $\{U_i\}$, we define the *cost-aware L2 norm* (CAL2N) in the i-th round as

$$\mathbb{E}\left\{\int_{T_i}^{T_{i+1}} |S_i(t)|^2 \, dt\right\} + c_0, \quad (2)$$

where $S_i(t)$ is defined in (1) and we use its L2 norm to characterize how far it has drifted away from 0. The constant $c_0 \geq 0$ is the *cost of reset* in the end of the round.

Our goal is to minimize the *long-term average* CAL2N:

$$\beta^*_{\text{CAL2N}} \triangleq \min_{\{X_i, M_i, U_i\}} \limsup_{n\to\infty} \frac{\sum_{i=1}^{n} \left(\mathbb{E}\left\{\int_{T_i}^{T_{i+1}} |S_i(t)|^2 \, dt\right\} + c_0\right)}{\sum_{i=1}^{n} \mathbb{E}\{T_{i+1} - T_i\}}. \quad (3)$$

To simplify (3), we notice that the optimization problem is a Markov decision problem with i.i.d.[3] Q_i, Y_i and V_i. As a result, it is sufficient to first find the optimal policy for the *single-round* optimization problem, assuming both the sensor and controller have access to some common randomness.[4] We can then apply the obtained solution to every round. The equivalent single-round optimization problem then becomes

$$\beta^*_{\text{CAL2N}} = \min_{X,M,U} \frac{\left(\mathbb{E}\left\{\int_0^{X+Y+U+V} |S(t)|^2\, dt\right\} + c_0\right)}{\mathbb{E}\left\{X+Y+U+V\right\}}. \tag{4}$$

C. The Setting in [15] as a Special Case

Our setting can be viewed as a strict generalization of [15] and [16]. Specifically, [15] solved[5] the Age-of-Information (AoI) minimization problem described as follows. Suppose there is a queue at the transmitter and the transmitter knows the status of the queue completely with zero delay. Every time the queue becomes empty (denoted as time T_i), the transmitter imposes a waiting time U_i and after that injects a packet to the queue, which takes V_i time to be serviced.

Suppose each packet is time-stamped and the AoI is defined as the current time minus the time stamp of the latest received packet. [15] observed that the AoI grows linearly over time and is intermittently reset to V_i at time $T_i + U_i + V_i$ since the time stamp of the latest arrival is $T_i + U_i$. Also see Fig. 2b. The goal is to minimize the long-term average AoI:

$$\beta^*_{\text{AoI}} \triangleq \min_{\{U_i\}} \limsup_{n\to\infty} \frac{\sum_{i=1}^{n}\mathbb{E}\left\{\int_{V_{i-1}}^{V_{i-1}+U_i+V_i} t dt\right\}}{\sum_{i=1}^{n}\mathbb{E}\left\{U_i+V_i\right\}} \tag{5}$$

$$= \min_{U_i} \frac{\mathbb{E}\left\{\int_{V_{i-1}}^{V_{i-1}+U_i+V_i} t dt\right\}}{\mathbb{E}\left\{U_i+V_i\right\}}, \tag{6}$$

where (6) is based on the equivalent single-round optimization.

We now show that (6) is a special case of our setting by (i) assuming $c_0 = 0$ and no forward delay $Y_i = 0$, (ii) hardwiring $X_i = M_i = 0$, i.e., forgoing the possibility of designing better X_i and M_i, and (iii) setting $Q_i = W(T_i) - W(T_{i-1}+U_{i-1})$. Define $T^{(i)} \triangleq \{T_j : j \le i\}$. We then have

$$\begin{aligned}
&\mathbb{E}\left\{|S_i(t)|^2 \middle| T^{(i)}\right\} \\
&= \mathbb{E}\left\{|W(t) - W(T_{i-1}+U_{i-1})|^2 \middle| T^{(i)}\right\} &(7)\\
&= t - (T_{i-1}+U_{i-1}) = V_{i-1} + (t - T_i), &(8)
\end{aligned}$$

where (7) follows from (1). To show (8), we note that knowing $T^{(i)}$ implies knowing $\{U_j : j \le i\}$ since the waiting time at the controller U_j is a function of $T^{(j)}$ (since $X_i = Y_i = M_i = 0$ in this specialized setting). This implies that the time instant $T_{i-1}+U_{i-1}$ is a deterministic value conditioning on $T^{(i)}$. By the strong Markov property we have (8). Finally, we can rewrite the numerator of (4) as

$$\mathbb{E}\left\{\int_{T_i}^{T_i+U_i+V_i} \mathbb{E}\left\{|S_i(t)|^2 \middle| T^{(i)}\right\} dt\right\} \tag{9}$$

$$= \mathbb{E}\left\{\int_{T_i}^{T_i+U_i+V_i} (V_{i-1} + t - T_i) dt\right\} \tag{10}$$

$$= \mathbb{E}\left\{\int_{V_{i-1}}^{V_{i-1}+U_i+V_i} t dt\right\}, \tag{11}$$

where (9) follows from Wald's equation [29] and the facts (i) T_i and U_i are deterministic values conditioning on $T^{(i)}$ (first proven in the previous paragraph) and (ii) V_i is independent of $T^{(i)}$ and $|S_i(t)|^2$. Eq. (10) follows from (8) and (11) follows from the change of variables. This proves that our objective function (4) collapses to the one for AoI minimization (6) once we set c_0, Y_i, X_i, M_i, and Q_i properly.

D. The Setting in [16] as a Special Case

Next, consider the remote estimation work in [16]. Suppose there is a queue at the transmitter and the transmitter knows the status of the queue instantaneously. Every time the queue becomes empty (denoted as time T_i), the transmitter continuously monitors the system state $W(T_i + t)$, which follows a Wiener process. After some waiting time $X_i \ge 0$, it sends the latest value $W(T_i + X_i)$ to the receiver, which takes Y_i time to be serviced and delivered to the receiver. The receiver then uses the latest value $W(T_i + X_i)$ as a system state estimate $\hat{W}(t) = W(T_i + X_i)$ during the subsequent time interval $[T_{i+1}, T_{i+1} + X_{i+1} + Y_{i+1})$, where $T_{i+1} \triangleq T_i + X_i + Y_i$ is the time when the current update packet arrives at the receiver, and $T_{i+1} + X_{i+1} + Y_{i+1}$ is the time when the next update packet arrives. As a result, the estimation error $W(t) - \hat{W}(t)$ jumps to a new initial value $Q_{i+1} = W(T_{i+1}) - W(T_i + X_i)$ at time T_{i+1}. The evolution of the estimation error $W(t) - \hat{W}(t)$ is described in Fig. 2c. See [16] for more details.

The objective of [16] is the following single-round optimization problem:

$$\beta^*_{\text{MMSE}} \triangleq \min_{X_i} \frac{\mathbb{E}\left\{\int_{T_i}^{T_i+X_i+Y_i} \left(W(t) - \hat{W}(t)\right)^2 dt\right\}}{\mathbb{E}\left\{X_i+Y_i\right\}}. \tag{12}$$

[3] More precise requirements are: (i) $\{Y_i\}$ and $\{V_i\}$ are i.i.d. and independent from the rest of the system; (ii) For any i, Q_i is independent of the waiting times $\{X_j, U_j : j < i\}$ in the previous rounds.

[4] The common randomness enables us to convert the temporal average over many rounds to the probabilistic average over a single round. In this work we implicitly assume the availability of common randomness when discussing any single-round optimization problem.

[5] [15], [16] also contained results of more advanced settings, e.g., arbitrary penalty functions, non-i.i.d. noises, etc. In this work, we focus on their simpler i.i.d. setting with no maximum update frequency constraint, which captures the simplest and most fundamental findings of their results.

We now show that the above remote estimation problem is a special case of our setting by (i) hardwiring $c_0 = U_i = V_i = 0$, i.e., short-circuiting the controller and the backward delay, (ii) setting $Q_i = W(T_i) - W(T_{i-1} + X_{i-1})$, and (iii) setting the message $M_i = S_i(T_i + X_i)$, the latest system state.

We can then rewrite the numerator of (4) as

$$\mathbb{E}\left\{\int_{T_i}^{T_{i+1}} \left(W(t) - W(T_i) + Q_i\right)^2 dt\right\} \tag{13}$$

$$= \mathbb{E}\left\{\int_{T_i}^{T_{i+1}} \left(W(t) - \hat{W}(t)\right)^2 dt\right\}, \tag{14}$$

where (13) follows from (1); (14) follows from the definitions of Q_i in the previous paragraph and $\hat{W}(t)$ during the time interval $[T_{i-1}, T_i)$. Our objective function (4) thus collapses to the one for remote estimation (12) once we set c_0, U_i, V_i, Q_i, and M_i properly.

In some sense, the two important papers [15], [16] form a perfect pair, the former of which focuses on the controller action (assuming $X_i = Y_i = 0$) while the latter studies the sensor action (assuming $U_i = V_i = 0$). The main contribution of this work is to unify these two results under a single framework and study the optimal joint sensor/controller scheme while fully taking into account random delay in both the forward and backward directions.

III. Main Results — The Optimal Policy

For all the analytical results, we provide the sketches of the proofs in Sec. V and leave the complete proofs to the journal submission.

A. An Auxiliary Minimization Problem

Given the distributions of the i.i.d. $\{Y_i\}$, $\{V_i\}$, $\{Q_i\}$ and some constant $c_0 \geq 0$, for any $\beta \geq 0$ we define $p(\beta)$ as the optimal value of the following minimization problem:

$$p(\beta) \triangleq \min_{(X,M,U)} \mathbb{E}\left\{\int_0^{X+Y+U+V} |S(t)|^2 dt\right\} + c_0 - \beta \mathbb{E}\{X+Y+U+V\}, \tag{15}$$

where we drop the subscript i for simplicity. We then have

Proposition 1: The function $p(\beta)$ is (i) continuous and strictly decreasing, (ii) there exists a unique $\beta^* \in [0, \beta_{\text{max}}]$ such that $p(\beta^*) = 0$, where

$$\beta_{\text{max}} \triangleq \text{Var}\{Q\} + \frac{\mathbb{E}\{Y\}\mathbb{E}\{V\} + \frac{1}{2}\mathbb{E}\{Y^2 + V^2\} + c_0}{\mathbb{E}\{Y+V\}} \tag{16}$$

(iii) the unique solution β^* is identical to the β^*_{CAL2N} defined in (4), and (iv) The (X_i, M_i, U_i) scheme that attains $p(\beta^*)$ also achieves the β^*_{CAL2N} in (4).

By Proposition 1, the minimization problem (4) can be solved in the following steps: For any given β, we first find the optimal (X, M, U) that minimizes (15) and the corresponding $p(\beta)$ value. We then find the optimal $\beta^* = \beta^*_{\text{CAL2N}}$ by a bisection search over $[0, \beta_{\text{max}}]$.

B. Optimal Waiting Time at the Controller

Define $\overline{M}_i = (T_i, X_i, M_i)$ as the information received by the controller in the i-th round, which consists of the time stamps (T_i, X_i) and the payload M_i.

Proposition 2: The optimal waiting time U_i^* at the controller that minimizes (15) is as follows.

$$U_i^* = \max\left(\beta - \left(Y_i + \mathbb{E}\left\{(S_i(T_i+X_i))^2 \,\middle|\, \overline{M}^{(i)}\right\} + \mathbb{E}\{V_i\}\right), 0\right), \tag{17}$$

where $\overline{M}^{(i)} \triangleq \{\overline{M}_j : j \leq i\}$.

That is, the optimal controller is a *water-filling policy* that calculates the difference between β and $(Y_i + \mathbb{E}\{(S_i(T_i+X_i))^2 \,|\, \overline{M}^{(i)}\} + \mathbb{E}\{V_i\})$. If $U_i^* = 0$, then the controller transmits the reset command immediately. Otherwise, the controller will impose a positive waiting time U_i^*.

C. Optimal Waiting Time at the Sensor

We now discuss the optimal M_i^* and X_i^* at the sensor.

Proposition 3: The optimal message that minimizes (15) is $M_i^* = S_i(T_i + X_i)$, the latest state value at the time of transmission $T_i + X_i$.

This intuitive result follows from the fact that the system state is a strong Markov process. We now discuss the optimal waiting time X_i^*, assuming the optimal U_i^* in Proposition 2 and the optimalM_i^* in Proposition 3 are used. For any $s \in (-\infty, \infty)$, define two functions $g_\beta(s)$ and $h_\beta(s)$ by

$$g_\beta(s) \triangleq a_{s,4} \cdot s^4 + a_{s,2} \cdot s^2 + a_{s,0} + a_0 + c_0 \tag{18}$$

$$h_\beta(s) \triangleq g_\beta(s) - (\beta s^2 - \frac{1}{6}s^4) \tag{19}$$

where

$$a_{s,4} \triangleq -\frac{\mathbb{P}\left(s^2 + Y \leq \beta - \mathbb{E}\{V\}\right)}{2} \tag{20}$$

$$a_{s,2} \triangleq \mathbb{E}\{Y+V\} + \mathbb{E}\left\{\mathbb{1}_{\{s^2+Y \leq \beta - \mathbb{E}\{V\}\}}\left(\beta - \mathbb{E}\{V\} - Y\right)\right\} \tag{21}$$

$$a_{s,0} \triangleq -\frac{\mathbb{E}\left\{\mathbb{1}_{\{s^2+Y \leq \beta - \mathbb{E}\{V\}\}}\left(\beta - \mathbb{E}\{V\} - Y\right)^2\right\}}{2} \tag{22}$$

$$a_0 \triangleq -\beta\mathbb{E}\{Y+V\} + \mathbb{E}\{Y\}\mathbb{E}\{V\} + \frac{1}{2}\mathbb{E}\left\{Y^2 + V^2\right\}$$

and $\mathbb{1}_{\{.\}}$ is the indicator function. Note that $g_\beta(s)$ and $h_\beta(s)$ are not fourth-order polynomials since the coefficients $a_{s,4}$, $a_{s,2}$, and $a_{s,0}$ also depend on s.

Lemma 1: For any β and any distributions of Y and V, which can be discrete, continuous, or hybrid, both functions $g_\beta(s)$ and $h_\beta(s)$ are even and continuous. Furthermore, there exist two constants $b_2 > 0$ and b_0 such that $g_\beta(s) \geq b_2 s^2 + b_0$ and $h_\beta(s) \geq \frac{1}{6}s^4 + b_0$ for all $s \in (-\infty, \infty)$.

For any two functions f_1 and f_2, we say $f_1 \prec f_2$ if $f_1(s) \leq f_2(s), \forall s \in (-\infty, \infty)$. The *lower convex envelope* (also called *the convex hull*) of the function $h_\beta(s)$ is defined as

$$\text{Cnvx}\big(h_\beta(s)\big) = \sup\{f(s) : f \text{ is convex}, f \prec h_\beta\}. \tag{23}$$

Lemma 2: For any $\beta \geq 0$, the lower convex envelope $\mathrm{Cnvx}\big(h_\beta(s)\big)$ is finite for all $s \in (-\infty, \infty)$.

We are now ready to describe the optimal waiting time X_i^* at the sensor.

Proposition 4: For any given β, the optimal X_i^* that minimizes (15) is the hitting time:

$$X_i^* = \inf\{t \geq 0 : S_i(T_i + t) \in \mathcal{S}_{\text{tx},\beta}\} \tag{24}$$

where the set $\mathcal{S}_{\text{tx},\beta}$, called *the transmission set*, is the collection of all state values s satisfying

$$\mathrm{Cnvx}\left(h_\beta(s)\right) = h_\beta(s), \tag{25}$$

i.e., the set of s whose corresponding values of the convex hull function are equal to those of the original function $h_\beta(s)$.

To be specific, the optimal sensor will hold off its transmission until the first time the state value $S_i(T_i+t)$ hits the transmission set $\mathcal{S}_{\text{tx},\beta}$. Fig. 3 illustrates a (piecewise) even function $h_\beta(s)$, which contains 5 pieces with the corresponding second-order derivatives being $+-+-+$ if we scan the s values from $-\infty$ to ∞. See the alternation between convexity and concavity for different values of s in Fig. 3. The convex hull function $\mathrm{Cnvx}\left(h_\beta(s)\right)$ is also drawn. One can see that in this example, $\mathrm{Cnvx}\left(h_\beta(s)\right) = h_\beta(s)$ iff $|s| \geq \gamma$ for some threshold γ. As a result, $\mathcal{S}_{\text{tx},\beta} = \{s : |s| \geq \gamma\}$. The optimal X_i^* is thus the first time when $|S_i(t)|$ hits γ.

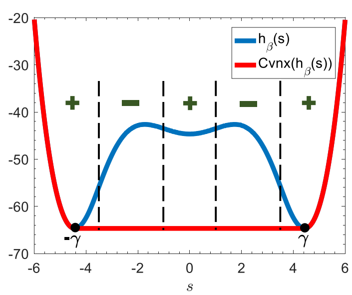


Fig. 3: An example of $h_\beta(s)$ and $\mathrm{Cnvx}(h_\beta(s))$.

D. Finding the Optimal β^*

Thus far we have described the optimal U_i^*, M_i^*, and X_i^* that minimizes (15). We now derive the minimum value $p(\beta)$.

Recall that Q_i describes the random state value in the beginning of the i-th round (right after reset). We then have

Proposition 5: Define $\phi(\beta, s)$ as

$$\phi(\beta, s) \triangleq \mathrm{Cnvx}\left(h_\beta(s)\right) + \beta s^2 - \frac{1}{6}s^4. \tag{26}$$

The optimal $p(\beta)$ in (15) can be computed by

$$p(\beta) = \mathbb{E}_Q\{\phi(\beta, Q)\}, \tag{27}$$

i.e., we first assume the s value in $\phi(\beta, s)$ is randomly distributed with distribution Q and then evaluate $p(\beta)$ by finding the expectation in (27).

We now summarize the detailed steps of finding the jointly optimal sensor/controller policy and the corresponding β^*_{CAL2N}.

Step 1: For any β, compute the functions $g_\beta(s)$, $h_\beta(s)$, and $\phi(\beta, s)$ by (18), (19), and (26), and compute $p(\beta)$ by (27).

Step 2: Repeatedly use Step 1 and the bisection search over $\beta \in [0, \beta_{\max}]$ to find the unique β^* satisfying $p(\beta^*) = 0$.

Step 3: Substitute $\beta = \beta^*$ in Secs. III-B and III-C to derive the respective optimal policies for the controller and the sensor.

We note that the bisection steps, i.e., Steps 2 and 3, also appear in [15], [16] and thus do not incur additional complexity. For some delay distributions Y_i and V_i, say, exponential, it is possible to derive closed-form expressions of g_β, h_β, and $\phi(\beta, \cdot)$ by calculus. For arbitrary Y_i and V_i distributions, we can compute g_β and h_β by quantizing the continuous s values into discrete points. Then we can use existing *linear-time* algorithms, e.g., [30], [31], to compute the convex hull $\mathrm{Cnvx}(h_\beta)$. The expectation step in (27) can subsequently be computed in linear time. Overall, the complexity of our algorithm is identical to [15], [16], all being linear-time in terms of the number of quantization points.

IV. Further Examination of The Optimal Policy

In this section, we prove some properties of the optimal joint sensor/controller scheme.

Lemma 3: $\mathcal{S}_{\text{tx},\beta^*}$ is symmetric over $s = 0$, i.e., for any $s \in (-\infty, \infty)$, $s \in \mathcal{S}_{\text{tx},\beta^*}$ if and only if $(-s) \in \mathcal{S}_{\text{tx},\beta^*}$.

Proof: This lemma follows directly from Lemma 1 and the definition of $\mathcal{S}_{\text{tx},\beta}$ (Proposition 4). ■

Define $\mathcal{S}^c_{\text{tx},\beta^*} \triangleq (-\infty, \infty)\backslash\mathcal{S}_{\text{tx},\beta^*}$ as the complement of $\mathcal{S}_{\text{tx},\beta^*}$. We then have

Lemma 4: $\mathcal{S}^c_{\text{tx},\beta^*}$ must be a collection of disjoint open intervals (l_i, r_i), namely,

$$\mathcal{S}^c_{\text{tx},\beta^*} = \bigcup_{i=1}^{\mu}(l_i, r_i) \tag{28}$$

where μ is the total number of open intervals and $\{(l_i, r_i) : i\}$ satisfies $-\infty < l_i < r_i < \infty$ for all $i \in [1, \mu]$.

Lemmas 3 and 4 imply that if $\mu = 1$, then $\mathcal{S}_{\text{tx},\beta^*} = \{s : |s| \geq \gamma\}$ for some $\gamma > 0$, which is termed a *threshold policy* in [16]. Similarly, if $\mu = 0$, then $\mathcal{S}_{\text{tx},\beta^*} = (-\infty, \infty)$ and the optimal policy is a zero-wait policy. In the sequel, we examine the value of μ, calculated by (28), for various scenarios.

A. Deterministic Forward Transmission Delay Y_i: $\mu \leq 1$

Proposition 6: If there exists a constant y_0 such that $P(Y_i = y_0) = 1$, then we always have $\mu \leq 1$ and $P(U_i^* = 0) = 1$.

In other words, with deterministic forward transmission delay Y_i, the optimal sensor policy is either a zero-wait policy ($\mu = 0$) or a threshold policy ($\mu = 1$), and the optimal controller strategy is always a zero-wait policy.

B. Exponential Forward Transmission Delay Y_i: $\mu \leq 2$

The delay of a communication channel is often modeled as a queue [14], and one of the most popular models is the exponential service time. We then have

Proposition 7: If $Y_i \sim \text{Exp}(\lambda_Y)$ is exponentially distributed with service rate $\lambda_Y > 0$, then we always have $\mu \leq 2$.

If we choose $Y \sim \text{Exp}(\lambda_Y)$ with $\lambda_Y = 0.2$, $V \sim \text{Exp}(\lambda_V)$ with $\lambda_V = 6$, $c_0 = 20$, and $Q \sim \mathcal{N}(0, \sigma^2)$ with $\sigma = 0.125$, then we can numerically compute $\beta^* = 7.24$ using the 3 steps in Sec. III-D. The resulting $\mathcal{S}_{\text{tx},\beta^*}$ indeed has $\mu = 2$. The upper bound $\mu \leq 2$ in Proposition 7 is thus tight.

We call the $\mu = 2$ policy an *interval policy*. The reason is that with $\mu = 2$ the transmission set $\mathcal{S}_{\text{tx},\beta^*}$ is of the form

$$\mathcal{S}_{\text{tx},\beta^*} = \{s : |s| \leq l \text{ or } r \leq |s|\} \tag{29}$$

for a pair of $0 < l < r < \infty$. That is, with a jointly optimal sensor/controller design, the optimal sensor should transmit its message $M_i^* = S_i(T_i + X_i)$ when the system state is either too large $|s| \geq r$ or too small $|s| \leq l$. At the first glance, this strategy seems counterintuitive due to the following reason: Our goal is to minimize the average value of $|S_i(t)|^2$. Therefore, large $|s|$ is considered to be "bad" and small $|s|$ is considered to be "good". An intuitive strategy that follows the thinking of the seminal works in [15] and [16] is to hold off transmission (i.e., to wait) when the state is good (when $|s|$ is small) in order to prolong the duration of staying in a good state. Our results show that when there is delay in both ways, the sensor sometimes should transmit when the state becomes *too good* (when $|s| \leq l$).

The explanation of this surprising phenomenon is as follows. The goal of CAL2N minimization in (4) is for the sensor and the controller to jointly design their strategies and *one thus has to decide how to split the waiting time between the sensor and the controller.* A deeper look at these two nodes shows that each of them has its unique advantage and disadvantage. In particular, the sensor is able to observe the full system state $S_i(t)$ continuously and uses it to make its decision X_i. The controller cannot observe $S_i(t)$ directly, but instead can directly observe the realization of the random sensor-to-controller delay Y_i, a valuable piece of information known exclusively to the controller. Therefore, when the system state is very good, $|s|$ being small, there is a bigger chance that the controller will see a good expected system state[6] $Y_i + \mathbb{E}\{(S_i(T_i + X_i))^2 | \overline{M}^{(i)}\} + \mathbb{E}\{V_i\}$ in (17). As a result, the sensor should transmit so that the controller, which has the additional observation of Y_i, can make a better informed decision U_i^* to further extend the duration of staying in a good system state. *One of the main contributions of this work is to uncover this unexpected sensor/controller coordination that is critical to achieving the optimal performance for systems with two-way random delay.*

The above discussion also explains the intuition of Proposition 6. With deterministic Y_i, the controller has a strictly inferior set of information since the observed Y_i is a constant. Hence, all the waiting time should be allocated to the sensor, i.e., $\mathbb{P}(U_i^* = 0) = 1$, and the sensor transmits if and only if the system state is bad (either a zero-wait or a threshold policy $|s| \geq \gamma$), which corresponds to $\mu \leq 1$.

[6]Since the optimal $M_i^* = S_i(T_i + X_i)$ is used, the term $\mathbb{E}\{(S_i(T_i + X_i))^2 | \overline{M}^{(i)}\} = (S_i(T_i + X_i))^2 = s^2$ is directly related to the value of $|s|$.

C. A Special Case of $\mu = 6$

The coordination between the sensor and the controller can sometimes be very subtle and beyond the high-level intuition discussed previously. Consider the following example.

Example 1: Consider the distribution of Y being

$$\begin{cases} \mathbb{P}(Y = 6) = 0.35, & \mathbb{P}(Y = 45) = 0.06, \\ \mathbb{P}(Y = 51) = 0.08, & \mathbb{P}(Y = 53) = 0.08, \\ \mathbb{P}(Y = 54) = 0.23, & \mathbb{P}(Y = 90) = 0.2. \end{cases} \tag{30}$$

$\mathbb{P}(V = 20) = 1$, $c_0 = 45$, and the initial random variable $Q \sim \mathcal{N}(0, \sigma^2)$ with $\sigma = 6.6$. We can numerically compute $\beta^* = 80.05$ using the 3 steps in Sec. III-D, and the resulting $\mathcal{S}_{\text{tx},\beta^*}$ has $\mu = 6$.

The reason of having a highly fractured transmission set ($\mathcal{S}_{\text{tx},\beta^*}^{\text{c}}$ containing $\mu = 6$ intervals) is due to the delicate probabilistic balance between the benefits of observing $S_i(t)$ at the sensor and observing Y_i at the controller.

V. High-level Sketches Of The Proofs

Our main results include the auxiliary minimization problem in Proposition 1, the optimal joint sensor/controller scheme in Propositions 2 to 5, and the discussion of the value of μ in Propositions 6 and 7. This extensive amount of analytical investigation is necessary in order to carefully present and analyze a new framework that unifies two recent important developments [15], [16]. In this section we present the high-level sketches of the corresponding proofs.

Proposition 1: $p(\beta)$ is continuous since we can use the available common randomness to *smoothen* the curve. $p(\beta)$ is strictly decreasing due to the assumption $\mathbb{E}\{Y + V\} > 0$ and its definition (15). The properties related to the root of $p(\beta)$ follow closely the derivation in [15], [16].

Propositions 2 to 5: The proofs of the optimal policy are based on casting the optimal joint sensor/controller policy for the Wiener process as the limiting case of the optimal dynamic programming (DP) solution over a binary random walk. Specifically, let $B(i)$ be a symmetric binary random walk such that $B(i+1) = B(i) + (1 - 2b(i))$, where $\{b(i) \in \{0,1\} : i\}$ are i.i.d. Bernoulli random variables with $p = 0.5$. We can use the discrete-time $B(i)$ to construct a continuous-time process $W^\delta(t)$:

$$W^\delta(t) = \sqrt{\delta} \cdot B(\left\lfloor \frac{t}{\delta} \right\rfloor) \tag{31}$$

where $\lfloor \cdot \rfloor$ is the floor function. It is well known that $W^\delta(t)$ converges to a Wiener process $W(t)$ in distribution when $\delta \to 0$. As a result, for any fixed $\delta > 0$ we first analyze the optimal joint sensor/controller scheme that minimizes (15), while assuming the underlying state value process is $W^\delta(t)$ (instead of $W(t)$). We then carefully characterize the limiting case $\delta \to 0$ and derive the results in Propositions 2 to 5.

Since $W^\delta(t)$ is based on the discrete-time $B(i)$, the corresponding problem of minimizing (15) is a discrete-time

Markov decision process (MDP) of an infinite horizon and an undiscounted objective function. To analyze its performance, we choose a large (but finite) horizon H, perform backward induction on the finite-horizon problem, solve for the optimal policy, and eventually let $H \to \infty$.

We now describe the backward induction under a given $H < \infty$. Since we have to jointly coordinate the sensor and the controller, it actually consists of two DP sub-problems where the first DP sub-problem corresponds to the controller decision U_i and the second DP sub-problem corresponds to the sensor decision X_i. We assume both DP problems are of the finite horizon H. We first solve for the optimal controller decision U_i under the discrete model $W^\delta(t)$ and the finite horizon H.

To model the sensor/controller coordination, the objective function of the first DP sub-problem is used to compute the "initial values" of the backward induction of the second DP sub-problem, i.e., the value function at the last (the H-th) time instant. After that, another round of backward induction is made to compute the optimal decision X_i at the sensor. The reason is that the optimal decision at the sensor must account for the optimal response from the controller, and thus we have to deal with two sequentially concatenated DP sub-problems.

Once we have fully solved the concatenated DP sub-problems, we compute the limits twice, i.e., first letting $H \to \infty$ and then letting $\delta \to 0$, to reach the final results. In fact, the $g_\beta(s)$ in (18) is the value function of the second (the senor side) DP at the end of the horizon after taking the double limits, and the $\phi(\beta, s)$ in (26) is the corresponding value function at the beginning of the horizon. The sequential steps of converting the $g_\beta(s)$ to the $\phi(\beta, s)$, especially the underlying convex-hull operations, capture the backward induction computation from the input $g_\beta(s)$ to the output $\phi(\beta, s)$.

Propositions 6 and 7: By specializing our general propositions to either deterministic Y_i or exponential Y_i, the expression of the final value function $\phi(\beta, s)$ can be explicitly found. We then carefully analyze the underlying convex-hull operations to derive the final results.

VI. Suboptimal Uncoordinated Policies

In this section we characterize two suboptimal schemes that do not exhibit any coordination between the sensor and controller: the optimal No-Wait-At-Sensor (NWAS) scheme and the optimal No-Wait-At-Controller (NWAC) scheme.

A. The Optimal No-Wait-At-Sensor (NWAS) Scheme

The best NWAS is optimized over M_i and U_i while hardwiring $X_i = 0$. A close look at the discussion in Sec. III shows that Propositions 2 and 3 can again be used to characterize the U_i^* and M_i^* of the NWAS scheme under any given β. We now describe how to find the optimal β^* of NWAS.

Proposition 8: For any $\beta \geq 0$, define $\phi_{\text{NWAS}}(\beta, s)$ and $p_{\text{NWAS}}(\beta)$ as

$$\phi_{\text{NWAS}}(\beta, s) \triangleq g_\beta(s) \tag{32}$$

$$p_{\text{NWAS}}(\beta) \triangleq \mathbb{E}_Q\{\phi_{\text{NWAS}}(\beta, Q)\} \tag{33}$$

where g_β is first defined in (18); Q is the initialized state random variable (after reset). The optimal β^* of the NWAS scheme is the unique root of $p_{\text{NWAS}}(\beta) = 0$, which can found by a numeric bisection search.

The main intuition of Proposition 8 is that since $X_i = 0$, the "horizon" of the second DP sub-problem is hardwired to 0, see Sec. V. Therefore, the $g_\beta(s)$ at the end of the horizon is the same as the $\phi_{\text{NWAS}}(\beta, s)$ at the beginning of the horizon. (32) essentially skips the intermediate steps that computes $\phi(\beta, s)$ in (26) from $g_\beta(s)$. (33) is in parallel to (27).

B. The Optimal No-Wait-At-Controller (NWAC) Scheme

The best NWAC is optimized over M_i and X_i while hardwiring $U_i = 0$, for which Proposition 3 can again be used to characterize the M_i^* of the NWAC scheme. We now describe how to find the optimal X_i^* under any given β and how to find the β^* of NWAC.

For any fixed β, the optimal waiting time X_i^* policy in Proposition 4 assumes the use of optimal U_i^*. Since we hardwire $U_i = 0$, we need to make some necessary changes. Particularly, we define a new $g_{\text{NWAC},\beta}(s)$ by

$$\begin{aligned} g_{\text{NWAC},\beta}(s) \triangleq & \mathbb{E}\{Y+V\} s^2 + \mathbb{E}\{Y\}\{V\} + \frac{1}{2}\mathbb{E}\{Y^2+V^2\} \\ & - \beta\mathbb{E}\{Y+V\} + c_0. \end{aligned} \tag{34}$$

Note that the $g_{\text{NWAC},\beta}(s)$ is a second-order polynomial of s since its coefficients do not depend on s.

By substituting $g_\beta(s) = g_{\text{NWAC},\beta}(s)$ in (18) and repeating the steps listed in Sec. III-C, we can find the X_i^* of the best NWAC policy. Specifically, we have

Proposition 9: For any given β, define $h_{\text{NWAC},\beta}(s) = g_{\text{NWAC},\beta}(s) - (\beta s^2 - \frac{1}{6}s^4)$. Since $g_{\text{NWAC},\beta}(s)$ has a nice form of being a second-order polynomial, by simple calculus one can verify that

$$\text{Cnvx}(h_{\text{NWAC},\beta}(s)) = \begin{cases} h_{\text{NWAC},\beta}(s) & \text{if } s^2 \geq \gamma_{\text{NWAC}} \\ h_{\text{NWAC},\beta}(\sqrt{\gamma_{\text{NWAC}}}) & \text{if } s^2 < \gamma_{\text{NWAC}}. \end{cases} \tag{35}$$

where $\gamma_{\text{NWAC}} \triangleq \max(3(\beta - \mathbb{E}\{Y+V\}), 0)$ is a constant threshold. As a result, the optimal transmission set of the NWAC scheme can be explicitly expressed as $\mathcal{S}_{\text{tx,NWAC},\beta} = \{s : s^2 \geq \gamma_{\text{NWAC}}\}$.

Similar to the discussion in Sec. III-C, the optimal β^* of NWAC is characterized by:

Proposition 10: Define $\phi_{\text{NWAC}}(\beta, s) \triangleq \text{Cnvx}(h_{\text{NWAC},\beta}(s)) + (\beta s^2 - \frac{1}{6}s^4)$ and $p_{\text{NWAC}}(\beta) \triangleq \mathbb{E}_Q\{\phi_{\text{NWAC}}(\beta, Q)\}$. The optimal β^* of NWAC is the unique root of $p_{\text{NWAC}}(\beta) = 0$.

In Sec. II-D, we have shown that the remote estimation problem is a special case of our setting with $U_i = 0$ and $M_i = S(T_i + X_i)$. As a result, the optimal remote estimation scheme can be viewed as an instance of NWAC ($X_i = 0$) scheme.[7] We then have the following corollary.

[7] The optimal AoI minimization scheme [15] is *not* an instance of the optimal NWAS scheme in Sec. VI-A since the former uses a suboptimal $M_i = 0$, see Sec. II-C. Conversely, the optimal NWAS scheme uses the optimal $M_i = S(T_i + X_i)$. As a result, the performance of the optimal AoI minimization scheme is no better than the optimal NWAS.

Corollary 1: When choosing the parameters c_0, U_i, V_i, and Q_i properly according to Sec. II-D, the optimal NWAC scheme described in this subsection indeed collapses to the optimal remote estimation scheme in [16, Theorem 1].

VII. Simulation Results

We compare the performance of our jointly optimal sensor/controller policy and five other important alternatives.

(i) Zero-wait [14]: Let $X_i = U_i = 0$, which is commonly known as the work-conserving policy in queueing theory.

(ii) Optimal No-Wait-At-Sensor (NWAS): We hardwire $X_i = 0$ and optimize over the message M_i and the controller waiting time U_i. See the discussion in Sec. VI-A.

(iii) Optimal No-Wait-At-Controller (NWAC): We hardwire $U_i = 0$ and optimize only over (X_i, M_i). See Sec. VI-B.

(iv) AoI-minimization scheme (AoI-min) [15]: As discussed in footnote 7 in Sec. VI-B, it is related to the NWAS scheme. The differences are (i) It falsely assumes the forward delay $Y_i = 0$ even though the actual Y_i is non-zero; (ii) It employs the suboptimal message $M_i = 0$ instead of the optimal $M_i^* = S_i(T_i + X_i)$, and (iii) It hardwires $X_i = 0$ and optimizes the U_i under the suboptimal assumptions (i) and (ii). We are interested in measuring the performance loss (compared to the optimal NWAS) due to these two suboptimal decisions.

(v) Remote-estimation scheme (RE) [16]: As discussed in Sec. VI-B, it is an instance of NWAC schemes. The difference between the RE and the optimal NWAC schemes is that the former falsely assumes the backward delay $V_i = 0$ even though the actual V_i is non-zero.

Due to space limits, we only report the numerical results for exponential forward and backward transmission delays. The initial value Q is Gaussian with zero mean and variance σ^2. The results are presented in Fig. 4.

We first notice that the larger the σ value, the wider the range of the initial value Q, which models the case of less accurate reset/control. Hence, as shown in Fig. 4a, the CAL2N of all 5 schemes increases as σ goes up.

A more interesting comparison is to calculate the ratio of the CAL2N of any scheme over that of our scheme, i.e., the normalized CAL2N plotted in Fig. 4b. Indeed, the normalized CAL2N of any scheme is always $\geq 100\%$ since our scheme is provably optimal. As expected, the AoI-min scheme is always worse (i.e., having higher CAL2N) than the optimal NWAS scheme. Similarly, the RE scheme is always worse than the optimal NWAC.

In Fig. 4b we also notice that when the reset is accurate (small σ), the performance of the optimal NWAS is identical to that given by the optimal solution. On the other hand, when the reset is loose (large σ), the performance of the optimal NWAC is identical to that of the optimal scheme. In either case, our algorithm optimally splits the waiting time between the sensor and the controller and always attains the best performance. In Figs. 4c to 4f, we varied the parameters λ_Y, λ_V, and c_0, and similar trends can be observed: Each of the 5 alternatives excels in some scenarios but performs poorly in the others, while our scheme always achieves the optimal performance.

In Secs. VI-A and VI-B, we design the best NWAS and NWAC schemes that correctly take into account the 2-way delay model. In contrast, the best existing schemes, i.e., the AoI-min and the RE schemes [15], [16], only consider the delay in one direction and completely ignore the delay in the other direction. The performance of the AoI-min or the RE schemes is thus much further away from optimality.

We also note that the effects of considering only 1-way delay (i.e., the AoI-min and RE schemes) and ignoring the delay in the other way is detrimental that in many cases they perform much worse than the naive zero-wait solution.

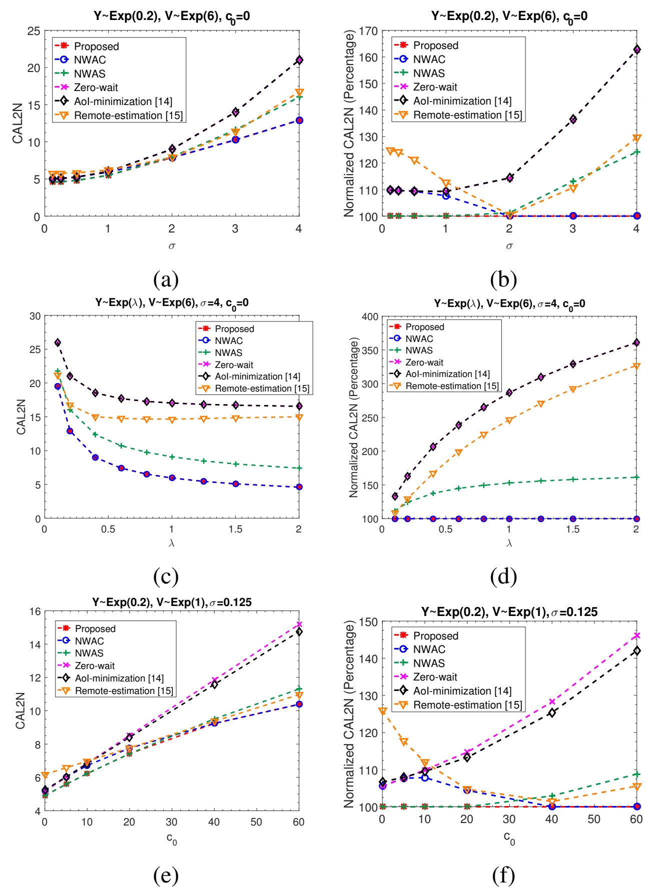


Fig. 4: Long-term average CAL2N for various settings. Those on the left are on absolute scale and those on the right are normalized with respect to the CAL2N of the optimal policy.

VIII. Conclusion

We have proposed a new Wiener-process-based framework and characterized the corresponding optimal policy, which unifies AoI minimization and remote estimation, two recent important results that have spawned substantial interests in the literature. The consideration of the two-way delay model and joint sensor-&-controller design has successfully addressed the double time asynchrony of the practical systems and represents a significant improvement over the existing results based on idealized zero-delay acknowledgement feedback.

References

[1] L. M. Huyett, E. Dassau, H. C. Zisser, and F. J. Doyle III, "Glucose sensor dynamics and the artificial pancreas: the impact of lag on sensor measurement and controller performance," *IEEE Control Systems Magazine*, vol. 38, no. 1, pp. 30–46, 2018.

[2] A. Javed, H. Larijani, A. Ahmadinia, R. Emmanuel, M. Mannion, and D. Gibson, "Design and implementation of a cloud enabled random neural network-based decentralized smart controller with intelligent sensor nodes for hvac," *IEEE Internet of Things Journal*, vol. 4, no. 2, pp. 393–403, 2016.

[3] T. Li, K. Keahey, K. Wang, D. Zhao, and I. Raicu, "A dynamically scalable cloud data infrastructure for sensor networks," in *Proceedings of the 6th Workshop on Scientific Cloud Computing*. ACM, 2015, pp. 25–28.

[4] E. Yanmaz, S. Yahyanejad, B. Rinner, H. Hellwagner, and C. Bettstetter, "Drone networks: Communications, coordination, and sensing," *Ad Hoc Networks*, vol. 68, pp. 1–15, 2018.

[5] Y. Zhang, M. Chen, N. Guizani, D. Wu, and V. C. Leung, "Sovcan: Safety-oriented vehicular controller area network," *IEEE Communications Magazine*, vol. 55, no. 8, pp. 94–99, 2017.

[6] X. Song and J. W.-S. Liu, "Performance of multiversion concurrency control algorithms in maintaining temporal consistency," in *Proceedings., Fourteenth Annual International Computer Software and Applications Conference*. IEEE, 1990, pp. 132–139.

[7] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *2012 Proceedings IEEE INFOCOM*. IEEE, 2012, pp. 2731–2735.

[8] C. Kam, S. Kompella, and A. Ephremides, "Age of information under random updates," in *2013 IEEE International Symposium on Information Theory*. IEEE, 2013, pp. 66–70.

[9] C. Kam, S. Kompella, G. D. Nguyen, and A. Ephremides, "Effect of message transmission path diversity on status age," *IEEE Transactions on Information Theory*, vol. 62, no. 3, pp. 1360–1374, 2015.

[10] M. Costa, M. Codreanu, and A. Ephremides, "Age of information with packet management," in *2014 IEEE International Symposium on Information Theory*. IEEE, 2014, pp. 1583–1587.

[11] L. Huang and E. Modiano, "Optimizing age-of-information in a multi-class queueing system," in *2015 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2015, pp. 1681–1685.

[12] B. T. Bacinoglu, E. T. Ceran, and E. Uysal-Biyikoglu, "Age of information under energy replenishment constraints," in *2015 Information Theory and Applications Workshop (ITA)*. IEEE, 2015, pp. 25–31.

[13] R. D. Yates, "Lazy is timely: Status updates by an energy harvesting source," in *2015 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2015, pp. 3008–3012.

[14] L. Kleinrock, *Queueing systems, volume 2: Computer applications*. Wiley-Interscience, 1976.

[15] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, and N. B. Shroff, "Update or wait: How to keep your data fresh," *IEEE Transactions on Information Theory*, vol. 63, no. 11, pp. 7492–7508, 2017.

[16] Y. Sun, Y. Polyanskiy, and E. Uysal-Biyikoglu, "Remote estimation of the Wiener process over a channel with random delay," in *2017 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2017, pp. 321–325.

[17] K. J. Astrom and B. M. Bernhardsson, "Comparison of Riemann and Lebesgue sampling for first order stochastic systems," in *Proceedings of the 41st IEEE Conference on Decision and Control, 2002.*, vol. 2. IEEE, 2002, pp. 2011–2016.

[18] B. Hajek, K. Mitzel, and S. Yang, "Paging and registration in cellular networks: Jointly optimal policies and an iterative algorithm," in *IEEE INFOCOM 2003. Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No. 03CH37428)*, vol. 1. IEEE, 2003, pp. 524–532.

[19] X. Gao, E. Akyol, and T. Başar, "Optimal estimation with limited measurements and noisy communication," in *2015 54th IEEE Conference on Decision and Control (CDC)*. IEEE, 2015, pp. 1775–1780.

[20] K. Nar and T. Başar, "Sampling multidimensional Wiener processes," in *53rd IEEE Conference on Decision and Control*. IEEE, 2014, pp. 3426–3431.

[21] E. Najm and R. Nasser, "Age of information: The gamma awakening," in *2016 IEEE International Symposium on Information Theory (ISIT)*. Ieee, 2016, pp. 2574–2578.

[22] K. Chen and L. Huang, "Age-of-information in the presence of error," in *2016 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2016, pp. 2579–2583.

[23] R. Durrett, *Stochastic calculus: a practical introduction*. CRC press, 2018.

[24] R. M. Mazo, *Brownian motion: fluctuations, dynamics, and applications*. Oxford University Press on Demand, 2002.

[25] R. P. Anderson and D. Milutinović, "A stochastic approach to dubins vehicle tracking problems," *IEEE Transactions on Automatic Control*, vol. 59, no. 10, pp. 2801–2806, 2014.

[26] S. Chung, A. Hoffmann, S. Bader, C. Liu, B. Kay, L. Makowski, and L. Chen, "Biological sensors based on Brownian relaxation of magnetic nanoparticles," *Applied physics letters*, vol. 85, no. 14, pp. 2971–2973, 2004.

[27] X. Lin, G. Sharma, R. R. Mazumdar, and N. B. Shroff, "Degenerate delay-capacity tradeoffs in ad-hoc networks with Brownian mobility," *IEEE/ACM Transactions on Networking (TON)*, vol. 14, no. SI, pp. 2777–2784, 2006.

[28] D. Lee, B. E. Carpenter, and N. Brownlee, "Media streaming observations: Trends in UDP to TCP ratio," *International Journal on Advances in Systems and Measurements*, vol. 3, no. 3-4, 2010.

[29] A. Wald, "On cumulative sums of random variables," *The Annals of Mathematical Statistics*, vol. 15, no. 3, pp. 283–296, 1944.

[30] S. B. Tor and A. E. Middleditch, "Convex decomposition of simple polygons," *ACM Transactions on Graphics (TOG)*, vol. 3, no. 4, pp. 244–265, 1984.

[31] A. A. Melkman, "On-line construction of the convex hull of a simple polyline," *Information Processing Letters*, vol. 25, no. 1, pp. 11–12, 1987.